

**REMARKS**

This paper is submitted in reply to the Office Action dated April 11, 2005. A request for a one month extension of time is enclosed herewith, and therefore, the period for response extends up to and includes August 11, 2005. Reconsideration and allowance of all pending claims are respectfully requested.

In the Office Action, the Examiner maintained the rejection of claims 1, 9, 12, 24, 29-30, and 37-38 under 35 U.S.C. § 102 (e) as being anticipated by U.S. Patent No. 6,480,838 to Peterman. Furthermore, the Examiner rejected claims 2-8, 10, 13-14, 16-18, 20, 25-29, 31-33, and 35 under 35 U.S.C. § 103 (a) as being unpatentable over Peterman further in view of U.S. Patent No. 6,519,586 to Anick et al.; rejected claims 1-10, 12-18, 20, 24-33, 35, and 37-39 as being unpatentable over Anick et al.; rejected claims 23 and 39 under 35 U.S.C. § 103(a) as being unpatentable over Anick et al. in view of Peterman; rejected claim 11 under 35 U.S.C. § 103 (a) as being unpatentable over Peterman and further in view of U.S. Patent No. 6,453,312 to Goiffon et al., and also rejected claim 11 under 35 U.S.C. § 103 (a) as being unpatentable over Anick et al. and further in view of Goiffon et al.; rejected claim 15 under 35 U.S.C. § 103 (a) as being unpatentable over Peterman and Anick et al. and further in view of U.S. Patent No. 5,799,276 to Komissarchik et al., and also rejected claim 15 as being unpatentable over Anick et al. and further in view of Komissarchik et al.; rejected claims 19, 21-22, 34, and 36 under 35 U.S.C. § 103 (a) as being unpatentable over Peterman and Anick et al. and further in view of U.S. Patent No. 6,044,387 to Angiulo et al., and also rejected claim 19, 21-22, 34 and 36 as being unpatentable over Anick et al. and further in view of Angiulo et al.; and rejected claims 40-43 as being unpatentable over Angiulo et al. in view of Anick et al.

Applicants respectfully traverse the Examiner's rejections to the extent that they are maintained. Applicants have amended claims 1, 23-24, and 37, and Applicants respectfully submit that no new matter is being added by the above amendments, as the

Page 10 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

amendments are fully supported in the specification, drawings and claims as originally filed.

As an initial matter, Applicants wish to thank the Examiner for the consideration granted in the telephone interview conducted on July 28, 2005. In the interview, the rejections based upon Peterman and Anick were discussed, but no agreement was reached.

Now turning to the subject Office Action and specifically to the rejection of claim 1, this claim generally recites a computer-implemented method of analyzing linguistic terms. The method includes scanning a plurality of documents for variants of a linguistic term, and tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during scanning to determine an acceptable usage of the linguistic term.

Applicants have also amended claim 1 herein to clarify that each variant of the linguistic term found in each document is "of the type that is intentionally chosen by an author of such document." Support for this amendment may be found, for example, at page 2, lines 17-20 of the Application as filed, wherein it is disclosed that variants are "chosen by other users," and it is clear from this passage as well as from the remainder of the Application that variants of a linguistic term are intended usages of a term by the respective authors of the documents that use the term. Put another way, the authors of the documents that are scanned generally do not accidentally use the terms in the manner in which they appear in the documents, and to the extent that authors may ever do so, such usages would in fact be detrimental to the overall operation of Applicants' invention, since one significant goal of embodiments consistent with the invention is to attempt to determine an acceptable usage of a term based upon how other users have chosen to use that term.

Along these same lines, the Examiner argues, in rejecting claim 1 as being anticipated by Peterman, that Peterman discloses the tracking of variants. However, it is

Page 11 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
E:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

clear from a reading of Peterman that these variants are due to OCR scanning errors, and not due to any intentional or chosen usages of variants by authors of documents. The variants in Peterman are in fact misspellings of terms that occur when an imaged document is scanned in and converted to text (e.g., where the term "in" is mistakenly interpreted by an OCR engine as the letter "m"). Peterman discloses, for example, that "potential OCR-produced alternatives (errors) for each character and character string and each alternative's probability of occurrence are determined." (col. 4, lines 9-11).

Peterman also discloses that:

"[i]t is desirable to have a mechanism that allows a search engine to accurately locate electronic documents that have been created using OCR software. Preferably, such a mechanism will recognize errors that are typically produced by OCR software and account for errors having the highest probability of occurrence." (col. 2, lines 28-34) (*emphasis added*).

Peterman therefore does not disclose the tracking of occurrences of variants of a linguistic term to determine an acceptable usage of the linguistic term, where each variant is "of the type that is intentionally chosen by an author of such document," as is required by claim 1. Claim 1 is therefore novel over Peterman, and the rejection should be withdrawn.

Claim 1 is also non-obvious over Peterman, as there is no suggestion in the art to modify Peterman to be used to track occurrences of variants of a linguistic term that are intentionally chosen by authors of the documents within which those variants are found. Indeed, Peterman itself teaches away from such a modification, as Peterman discloses at col. 2, lines 5-27, that a prior art technique that uses fuzzy searching can undesirably result in too many hits because of the inclusion of alternate search terms that represent other legitimate terms. Peterman also states, at col. 3, lines 13-16, that its manner of creating search term variants (which is based upon the probability of an OCR error occurring for a particular character or string of characters in a term) is an improvement

Page 12 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\bm\166\Amendment and Response re 4-11-05 OA.wpd

over the art "because the variant words constructed in this manner are rarely legitimate words in the natural language of the database." It is evident that Peterman attempts to generate search terms that will not generate false matches with terms that are "legitimate," or intended to be used, by the authors of any searched documents.

Claim 1 was also rejected as being obvious in view of Anick. The Examiner argues at ¶18 of the subject Office Action that Anick discloses searching for and tracking the occurrence of "facets." However, Applicants are unsure as to whether the Examiner intended to maintain this rejection, as the Examiner states as follows on page 34 of the subject Office Action, in connection with the rejection of claim 23:

Applicants argue that Anick does not teach or suggest tracking relative occurrences of a plurality of variants of a linguistic term and determining an acceptable usage of the linguistic term . . . Examiner agrees. Peterman, in combination with Anick, teaches and suggests the argued feature.

It appears from this statement that the Examiner has taken the position that Anick, by itself, does not disclose or suggest this feature, which is also recited in claim 1, and with which Applicants heartily agree. Moreover, it also appears that the Examiner feels that Peterman discloses this feature, and is appropriately combined with Anick. Nonetheless, as Applicants have established above, Peterman does not in fact disclose or suggest this feature, particularly given claim 1 has been amended to clarify that each recited variant is "of the type that is intentionally chosen by an author of such document."

As Applicants noted in the prior response, Anick discloses that new concepts are often expressed as concatenations of existing terms instead of new, single words. Anick discloses locating facets of a particular search term through lexical dispersion that measures the number of different words with which that particular word co-occurs within such syntactic contexts. The more different words that are identified, the more likely that

Page 13 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

the search term will represent a useful conceptual category relevant to the query. (see column 2, lines 40-55).

In addition, in Anick the facets are presented to a user in connection with the results of a search based upon the search term, and are used to reformulate a query to effectively narrow the results to focus on a particular conceptual category. As shown in Fig. 3B, for example, a query using the search term "cooking" results in 265 hits, along with the display of a number of facets that can be used to narrow the search. When a user selects one of the facets, e.g., "pecorino cheese," the search is focused to include only those documents that contain both "cooking" and "pecorino cheese."

Thus, the "facets" of Anick are not variants of a linguistic term. Indeed, even looking at the example of Fig. 3B, it is evident that the different facets relating to the word "cheese" are different terms altogether, not variations of a single term. Facets are more appropriately characterized by Anick as "key informational concepts" that can be identified within a document based on a query search term. The facets of Anick involve identifying relations between a search term and co-occurring words for the purpose of defining different contexts in which the search term might be used.

In contrast, the "variants of a linguistic term" recited in claim 1 relate to variations of that term, not to co-occurring words, and the purpose of identifying the variants in the claims is to determine an acceptable usage of the linguistic term, not to identify contextual domains for that linguistic term.

Applicants therefore respectfully submit that Anick, whether considered alone or in combination with Peterman, does not disclose or suggest the invention recited in claim 1. Moreover, given that Peterman expressly teaches away from the invention recited in claim 1, Applicants respectfully submit that one of ordinary skill in the art would not look to modify either of Peterman or Anick to arrive at Applicants' claimed invention. Claim 1 is therefore non-obvious over the prior art of record, and reconsideration and allowance

Page 14 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response rc 4-11-05 OA.wpd

of claim 1, as well as of claims 2-22 which depend therefrom, are therefore respectfully requested.

Next with respect to independent claim 23, this claim generally recites a method of analyzing linguistic terms. The method includes browsing a plurality of web sites on the Internet in response to user input, and concurrently with browsing the plurality of web sites, tracking relative occurrences of a plurality of variants of a linguistic term found in the plurality of web sites to determine an acceptable usage of the linguistic term.

Claim 23 has also been amended in a similar manner to claim 1, and now recites that each variant of the linguistic term found in each web site is "of the type that is intentionally chosen by an author of such web site."

Claim 23 is rejected based upon Anick in view of Peterman, and in connection with this rejection, the Examiner acknowledges that Anick does not disclose tracking relative occurrences of a plurality of variants of a linguistic term to determine an acceptable usage of a linguistic term. Instead, the Examiner relies on Peterman for allegedly disclosing this concept. However, as discussed above in connection with claim 1, Peterman not only does not disclose or suggest, but expressly teaches away from, tracking occurrences of variants of a linguistic term where those variants are of the type that are intentionally chosen. Therefore, Applicants submit that claim 23 is patentable over the combination of Anick and Peterman for the same reasons as presented above for claim 1. Reconsideration and allowance of claim 23 are therefore respectfully requested.

Next with respect to independent claims 24 and 37, each of these claims has been amended in a similar manner to claim 1, and now recite the concept of determining an acceptable usage of a linguistic term at least in part by tracking relative occurrences of a plurality of variants of the linguistic term, where those variants are of the type that are intentionally chosen by authors of scanned documents where those variants are found. As discussed above in connection with claim 1, this concept is not disclosed or suggested by either of Peterman or Anick, whether alone or in combination. Accordingly, claims 24

Page 15 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

and 37 are patentable over the prior art of record for the same reasons presented above for claim 1. Reconsideration and allowance of claims 24 and 37, as well as of claims 25-36 and 38 which depend therefrom, are therefore respectfully requested.

Next turning to the rejection of independent claim 39, this claim generally recites a program product that includes a document, and a computer-readable signal bearing medium bearing the document. The document includes a tag that identifies an acceptable variant of a linguistic term and a definition of the linguistic term.

In rejecting claim 39 based upon the combination of Anick and Peterman, the Examiner states only that "claim 39 is a program product of method claim 23." However, claim 39 recites a tag in a document that identifies "an acceptable variant of a linguistic term and a definition of the linguistic term," a feature that is not recited in claim 23.

Moreover, neither Anick nor Peterman discloses any document with an embedded tag that includes either a variant of a linguistic term or a definition of the linguistic term. Precisely how the Examiner considers the references to teach these concepts is unclear, as Applicants can not even find anything even arguably relevant to these concepts in either reference. Accordingly, Applicants submit that the Examiner has failed to provide the necessary basis to establish a prima facie case of obviousness under 35 U.S.C. §103. Reconsideration and withdrawal of the rejection of claim 39 are respectfully requested.

Next turning to the rejection of independent claim 40, this claim generally recites a method of spell checking a document. The method includes comparing terms in a first document against an electronic dictionary, and in response to determining during the comparison that a term from the document is not in the electronic dictionary, automatically scanning a plurality of documents from the Internet to identify at least one acceptable usage of the term.

In rejecting claim 40, the Examiner relies on Angiulo combined with Anick. Angiulo only discloses conventional spell checking, but applied to multiple documents, e.g., from a common web site. In fact, the Examiner admits the narrow teachings of

Page 16 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

Angiulo by noting that Angiulo does not disclose Applicants' claimed feature of "in response to determining during the comparison that a term from the document is not in the electronic dictionary, automatically scanning a plurality of documents from the Internet to identify at least one acceptable usage of the term."

Instead, the Examiner relies on Anick for allegedly disclosing this latter feature. However, as discussed above in connection with claim 1, Anick does not disclose or suggest the concept of scanning a plurality of documents to identify an acceptable usage of a term. Anick discloses performing a search based upon a search term. Anick also discloses generating facets for a corpus of documents to assist in focusing a search toward a particular contextual usage of a search term. However, Anick does not disclose or suggest scanning documents for the purpose of identifying an acceptable usage of a term, as required by claim 40. Applicants therefore respectfully submit that the combination of Angiulo and Anick fails to disclose or suggest each and every feature of claim 40.

Furthermore, even if Anick did disclose or suggest scanning documents for the purpose of identifying an acceptable usage of a term, neither Anick nor Angiulo discloses or suggests automatically triggering such a scan "in response to determining during [a] comparison that a term from [a] document is not in [an] electronic dictionary." Put another way, claim 40 covers a scenario where a document is spell checked against an electronic dictionary, and based upon not finding a term in the document in the dictionary, a scan is automatically initiated on a plurality of documents from the Internet to attempt to identify an acceptable usage of the term. In this manner, a user may be able to find out if a term that is not in an electronic dictionary is nonetheless acceptable based upon a search of potential usages of the term by others.

It appears that the Examiner has completely disregarded this triggering mechanism in analyzing claim 40. Angiulo discloses spell checking, but never discloses or suggests that any particular search or similar operation can be initiated automatically in response to not finding a particular term in a document. Conversely, Anick discloses

Page 17 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd



performing searching of a corpus of documents, but is completely silent with regard to spell checking. Given these unrelated teachings, one of ordinary skill in the art would not look to Anick to modify Angiulo in the manner suggested by the Examiner, and Applicants submit that the Examiner is relying on hindsight in proposing the combination of these references. There is simply no motivation in the art to combine the references to arrive at the unique functionality recited in claim 40.

Applicants therefore respectfully submit that claim 40 is patentable over the prior art of record, and reconsideration and allowance of claim 40, and of claim 41 which depends therefrom, are respectfully requested.

Next, turning to the rejection of independent claim 42, this claim generally recites a method of managing an electronic dictionary. The method includes detecting a spell definition tag within a document retrieved from the Internet that identifies an acceptable variant of a linguistic term, and in response to detecting the spell definition tag, automatically adding the acceptable variant of the linguistic term to an electronic dictionary.

Again, the Examiner relies on the combination of Angiulo and Anick; however, the Examiner does not indicate what, if any, relevance Anick has to the claim. Instead, the Examiner appears to completely rely on Angiulo, and in particular the conventional capability of a spell checker in scanning a document and being updated by a user with a spelling that is not found in an electronic dictionary.

The Examiner admits that Angiulo does not disclose a spell definition tag, and argues that it was well known to include tags for various reasons in a web document. With all due respect, however, Applicants assert that such an argument is replete with hindsight based reasoning, as absolutely no objective evidence of a motivation to incorporate a tag for this specific reason is provided by the Examiner.

Applicants submit that the fact that tags in web documents are known falls far short of disclosing or suggesting a specific type of tag that can be embedded in a

Page 18 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\km\166\Amendment and Response re 4-11-05 OA.wpd

document and used to automatically add a term to an electronic dictionary. Angiulo discloses at the most that a user, when presented with a term that is not in an electronic dictionary, can opt to manually add that term to the electronic dictionary. Angiulo does not even suggest that a term can be added automatically to an electronic dictionary in response to any operation other than a manual operation that is initiated by a user when an unrecognized term is found during a spell check, much less in response to detecting a particular type of tag that is embedded in a retrieved document.

The spell definition tag recited in claim 42 provides a unique and unexpected advantage over Angiulo and the other art of record, by potentially enabling the author of a document to automatically cause an electronic dictionary for the reader of a document to be updated with an acceptable variant of a term. Put another way, a document created by one individual can cause the electronic dictionary for another individual to automatically add a term. Moreover, it is important to note that the claimed operation need not occur when a spell check operation is being performed, but rather occurs in response to a tag being detected in a document that is retrieved from the Internet. There is no disclosure in Angiulo that even arguably suggests such functionality.

Given also that Anick does not even deal with electronic dictionaries, it appears that Anick adds nothing to the Examiner's rejection.

Applicants therefore respectfully submit that claim 42 is patentable over the prior art of record. Reconsideration and allowance of claim 42, as well as of claim 43 which depends therefrom, are therefore respectfully requested.

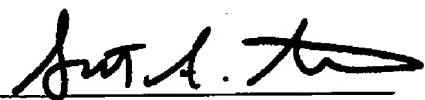
As a final matter, Applicants traverse the Examiner's rejections of the various dependent claims on the basis of the dependency of these claims upon the aforementioned independent claims. While a number of these claims recite additional features that are neither disclosed nor suggested by the prior art of record, Applicants will not address these claims separately at this time.

Page 19 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd

In summary, Applicants respectfully submit that all pending claims are novel and non-obvious over the prior art of record. Reconsideration and allowance of all pending claims are therefore respectfully requested. If the Examiner has any questions regarding the foregoing, or which might otherwise further this case onto allowance, the Examiner may contact the undersigned at (513) 241-2324. Moreover, if any other charges or credits are necessary to complete this communication, please apply them to Deposit Account 23-3000.

Respectfully submitted,

11 AUG 2005  
Date

  
Scott A. Stinebruner  
Reg. No. 38,323  
WOOD, HERRON & EVANS, L.L.P.  
2700 Carew Tower  
441 Vine Street  
Cincinnati, Ohio 45202  
Telephone: (513) 241-2324  
Facsimile: (513) 241-6234

Page 20 of 20  
Serial No. 09/751,574  
Amendment and Response dated August 11, 2005  
Reply to Office Action of April 11, 2005  
IBM Docket ROC920000191  
WH&E IBM/166  
K:\ibm\166\Amendment and Response re 4-11-05 OA.wpd